Impacts of interpolation and adaptive inflation when assimilating radar reflectivity using an ensemble adjustment Kalman filter

Thomas A. Jones[1]

[1.] Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, OK

Corresponding Author Address:
Dr. Thomas A. Jones
Cooperative Institute for Mesoscale Meteorological Studies
University of Oklahoma
and
NOAA/OAR/National Severe Storms Laboratory
120 David L. Boren Blvd.
Norman, OK 73072

Abstract

Various techniques have been used to successfully assimilate ground based radar reflectivity into numerical weather prediction models. Assimilating reflectivity coupled with Doppler radar radial velocity has led to significant advances in storm-scale modeling of convection that produces severe weather. Assimilation of radar reflectivity using an ensemble Kalman filter requires the computation of simulated reflectivity using the hydrometeor mixing ratios and number concentrations contained in the model state. In several previous studies, reflectivity has been first calculated on the model grid using the model microphysics scheme and then included in the state vector, after which it is interpolated to the observation location. Alternatively, it is also possible to calculate reflectivity by first interpolating the state variables to the observation location and then subsequently computing reflectivity at the location. Since the reflectivity forward operator is nonlinear, the two methods are not identical. This research compares the differences between these two methods. Results from an idealized case indicate that calculating reflectivity in the forward operator when using adaptive inflation results in a storm with higher concentrations of rainwater mixing ratio, greater updraft velocities, and a cooler cold pool. Similar results were found in the real-data comparisons made using the 20 May 2013 central Oklahoma tornado event. In this case, 0-1 hour forecasts of precipitation were improved when first interpolating the microphysical state variables to the observation points and then computing the reflectivity. While both assimilation methods have their advantages and disadvantages, the selection of which one to use can have a large impact on the resulting model analyses and ensuing forecasts.

1. Introduction

Assimilating radar reflectivity into storm-scale numerical weather prediction (NWP) models has proven to be a challenging task, but has also shown to have great potential for improving severe weather forecasting (e.g. Gao et al. 1999; Dowell et al. 2004; Tong and Xue 2005; Tong and Xue 2008a,b; Aksoy et al. 2009, 2010; Yussouf and Stensrud 2010; Dowell et al. 2011; Gao and Stensrud 2012; Yussouf et al. 2013; Wheatley et al. 2015). Radar reflectivity is directly related to the cloud hydrometeor properties of precipitating clouds through the electromagnetic backscatters from a ~0.5-2 km^3 volume of air having a variety of particle sizes and densities (Doviak and Zrnic 1993). Often, mixed phase particles (water and ice) are present in a particular volume. On the other hand, storm-scale models produce reflectivity grids that are generated from a variety of cloud microphysical parameterizations. The computation of reflectivity from cloud hydrometeors is a function of the particular cloud microphysics scheme used within the model (Tong and Xue 2008a,b; Jung et al 2012; Yussouf et al. 2013). Even when using multi-moment microphysical schemes, significant biases and errors have been demonstrated relative to the observed reflectivity structures in a variety of convective storms (Dawson et al. 2012; Jung et al 2012; Yussouf et al. 2013).

When assimilating radar observations, the practical challenge for storm-scale data assimilation is to determine an accurate set of background error covariances that can be used to update the model's state from the set of available observations. Since storm-scale data assimilation is primarily a retrieval problem (rather than an analysis problem), the characteristics of the background state strongly controls the resulting analysis. The forecast model's inherent

biases and errors are difficult to determine due to a lack of knowledge of the internal state of convective storms, and therefore errors can significantly impact the resulting predictions. These issues are particularly true for reflectivity assimilation, as current state-of-the-art microphysics parameterizations, while much improved over the last twenty years, still contain many random errors and biases (Tong and Xue 2008a,b; Dawson et al. 2012). Complicating this is the nonlinear relationship between reflectivity and cloud hydrometeors leaving the forward operator sensitive to a particular scheme choice or scheme parameters chosen for the forecast model (Tong and Xue 2008a,b; Jung et al 2012; Yussouf et al. 2013).

Several different techniques have been applied to radar reflectivity data assimilation with varying degrees of success. Hu and Xue (2007) derived a relationship between composite reflectivity and atmospheric latent heating tendency producing a parameter that can be assimilated into a NWP model without having to consider the uncertainties between reflectivity and hydrometeor concentrations as a function of microphysics. More recent research has focused on directly assimilating radar reflectivity through its known relationships to the various hydrometeor properties. This approach allows for a much better 3-D representation of convection within storm-scale models than the latent heating method. Many of these studies use the ensemble adjustment Kalman filter (EAKF) approach (e.g. Anderson and Collins 2009) with high degrees of success (e.g. Tong and Xue 2008a,b, Aksoy et al. 2009, 2010, Dowell and Wicker 2009; Yussouf et al. 2013).

In principle, assimilating reflectivity using the EAKF approach should generate a realistic initial distribution of hydrometeors within the storm via the use of the background error covariances that are evolving and highly dependent on the current maturity and size of the convective storm. However, reflectivity biases in the model forecasts can have a strong negative impact on the analysis of state fields other than hydrometeors. Dowell and Wicker (2009) demonstrated that negative biases within a supercell's forward flank reflectivity region at low levels can lead to excessive cold pool deficits in the forward flank region of the storm. Since ensemble mean reflectivity often has a low bias in these regions, and because the correlation between reflectivity and potential temperature is often negative, the analysis continually adjusts the temperature field to be colder. These biases then lead to unrealistically cold temperatures in the storm's forward flank cold pool. The use of adaptive inflation (Anderson 2007, 2009) is needed to maintain ensemble spread, but can also exacerbate the biases in the non-linear forward operators, particularly when multiplicative inflation is used.

The work presented here explores in more detail the analysis and forecast impacts of various reflectivity assimilation methods used within the EAKF. In particular, we will demonstrate that the nonlinear reflectivity forward operator and the adaptive inflation, applied at different stages of the analysis cycle, can generate significant changes in the resulting ensemble state's analyses. We use the EAKF present in the Data Assimilation Research Testbed (DART) system in combination with the WRF-ARW model to first examine the issue using an idealized observing system simulation experiment (OSSE) and then show similar behavior in a real data case occurring on 20 May 2013. The results show that the analysis system is sensitive to whether the reflectivity is first computed on the model grid and then subsequently included as a component of the state vector (rather than being computed internally from a set of interpolated hydrometeors/number concentrations within the data assimilation system algorithms) and that due to the nonlinearity of the reflectivity operator, the use of adaptive inflation on the observational state's priors also impacts the results significantly. Hereafter, we will refer to the

temperature, hydrometeor mixing ratios, and number concentration variables needed to compute reflectivity from a particular microphysics scheme as the "microphysical variables".

The goal of this research is to identify the origin of the differences observed in the posterior analyses caused by the manner in which radar reflectivity is computed and used in the assimilation process. Following the Introduction, Section 2 provides an overview of the two radar reflectivity assimilation techniques compared as part of this research. Section 3 summarizes the EAKF and its relationship to radar reflectivity assimilation. Section 4 describes the idealized nature run and OSSE experiment design while Section 5 discusses the results of the idealized experiments. A discussion of the 20 May 2013 results follows in Section 6 with concluding remarks in Section 7.

2. Reflectivity Assimilation Techniques

The reflectivity factor (Z) is a measure of the radiation reflected back to a radar from a volume of scatterers, often precipitation hydrometeors (Rinehart 1997). Reflectivity is related to the characteristics of hydrometeors in a highly non-linear manner. It can be represented by the summation of the number of hydrometeors (N_i) with diameter (D_i) over all drop sizes within a sample volume (Eq. 1).

$$Z = \sum N_i D_i^6 \tag{1}$$

Values for radar reflectivity factor can range from (< 1.0 mm^6 m^{-3}) for light precipitation to $> 10^7$ mm^6 m^{-3} for heavy rain and hail. Due to the large range of values, reflectivity factor is commonly converted into logarithmic units (decibels of reflectivity or dBZ) using Eq (2).

$$dBZ = 10 \log_{10} Z \tag{2}$$

For the remainder of this research, the term reflectivity will refer to logarithmic reflectivity in units of dBZ.

When assimilating radar reflectivity, a forward operator is required to generate reflectivity from the model state. In the ideal case, this forward operator should use code consistent with the model's cloud microphysics scheme. Depending on which method is chosen, there are two places where the reflectivity forward operator can be applied within a "parallel" EAKF algorithm (Anderson and Collins 2007; Whitaker et al. 2008). First, the forward operator is applied on each ensemble member's grid and the gridded reflectivity field is included in the ensemble state vector, similar to other state variables such as potential temperature or vertical velocity. This "external" application of the forward operator (hereafter denoted as EFO) is applied prior to the assimilation step. The gridded reflectivity is linearly interpolated directly to reflectivity observation location to create the priors needed for the data assimilation. The parallel algorithm then extends the ensemble state vector by appending all of the observational priors, from which the analysis update can be computed. Generally, it is reflectivity in dBZ space that is interpolated, not the original reflectivity factor, Z. Given the highly nonlinear nature of Z, it is likely that interpolating one or the other will produce different results when assimilating radar

reflectivity observations. As a result, both methods are tested within this research using the EFO assimilation technique.

It is also possible to calculate reflectivity (dBZ) internally in the forward operator from the linearly interpolated hydrometeor state variables (Aksoy et al. 2009). This internal method, hereafter denoted as IFO, interpolates the individual hydrometeor variable in the state to the observation location after which the microphysics specific radar reflectivity formula is applied. The hydrometeor mixing ratio and number concentration variables are by nature nonlinear terms, which is problematic in a data assimilation system that often assumes linear properties. However, the reflectivity value calculated from the interpolated state is also more likely to be consistent with the hydrometeor characteristics at that location compared to reflectivity calculated using either Z or dBZ interpolation in the EFO method.

Also, since reflectivity is no longer included in the state vector, its overall size becomes smaller, reducing the overhead required to run the EAKF. Additional complications become apparent when attempting to assimilate other closely related data sets such as cloud water path (CWP) or infrared radiances from satellite data (Jones et al. 2013a,b, 2014, 2015). For satellite data assimilation, the common approach is to calculate parameters such as CWP and/or infrared radiances internally within the data assimilation system using interpolated state variables. When assimilating both radar and satellite data, it was found that using the EFO method coupled with satellite derived priors computed internally (i.e., using an IFO approach for the satellite forward operator) could lead to poor analysis and forecasts in certain conditions. When reflectivity was computed internally from interpolated microphysical variables (e.g. IFO), the model analyses and forecasts were often improved.

3. Parallel ensemble adjustment Kalman filter

To understand the differences between EFO and IFO, it is important to carefully describe the process by which these data are assimilated using the EAKF algorithm implemented in DART (Fig. 1). The prior state vector is first generated from the previous cycle of ensemble forecasts. The state vector generally includes prognostic variables such as pressure, temperature, moisture, hydrometeor characteristics, wind components and potentially diagnostic variables such as radar reflectivity (Table 1). Prognostic variables represent those that are modified during model integration and impact the forecast. Diagnostic variables are derived from prognostic variables and are not needed to advance the prognostic system forward in time. For example, the updated analysis of radar reflectivity from the EFO method has no direct impact on the model forecast, the field is simply replaced by the microphysics scheme's calculation at the next time step.

VARIABLE	UNITS	TYPE
U-WIND	ms^{-1}	P
V-WIND	ms^{-1}	P
W-WIND	ms^{-1}	P
GEOPOTENTIAL HEIGHT	m	P
POTENTIAL TEMP	K	P
PRESSURE	Pa	D
WATER VAPOR	$kg\,kg^{-1}$	P
CLOUD WATER	$kg\,kg^{-1}$	P
RAIN	$kg\,kg^{-1}$	P
ICE	$kg\,kg^{-1}$	P
SNOW	$kg\,kg^{-1}$	P
GRAUPEL	$kg\,kg^{-1}$	P
ICE NUMBER	Number	P
RAIN NUMBER	Number	P
DIABATIC HEATING	$K\,s^{-1}$	P
RADAR REFLECTIVITY	dBZ	D

Table 1. *List of model variables includes in EAKF state vector for the idealized experiments. All variables except radar reflectivity are considered prognostic (P) variables. Reflectivity is a diagnostic (D) variable that is calculated from prognostic variables in the model and has no impact on model forecasts. All IFO experiments remove radar reflectivity from the state.*
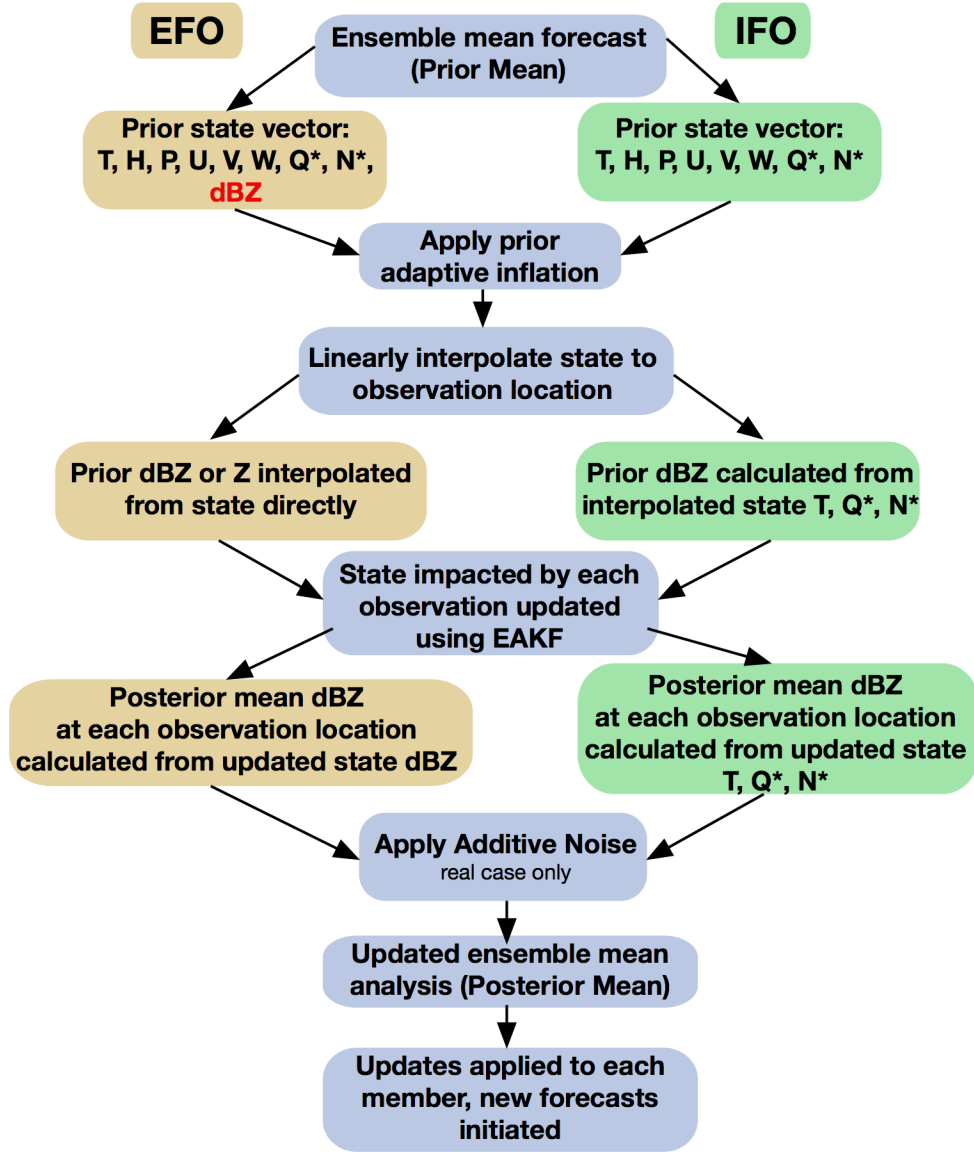
Figure 1. *Flow chart describing the steps required for assimilating radar reflectivity into the EAKF system using external (EFO) and internal (IFO) methods.*

Prior adaptive covariance inflation is applied to increase ensemble spread before data assimilation occurs (Anderson 2007, 2009). Inflation is required to maintain spread in the ensemble members to counteract the collapse of spread due to under-sampling of the true ensemble variance from repeated cycling. The DART system's adaptive inflation (AI) factor is a time and state variable-dependent and evolves as a function of time. It has dimensions M x N, where M is the number of grid points and N is the number of state variables. The inflation factor inflates the perturbations between the prior state and observations. The larger the magnitude of the perturbations, the larger the AI applied. In the EFO method, reflectivity would have its own AI field and this field is inflated prior to assimilation.

For the IFO, reflectivity is computed from the inflated hydrometeors mass and number, which means that each variable used to calculate reflectivity is inflated independently prior to the calculation of reflectivity. After the application of AI, the prior state vector is linearly

interpolated from model space (x,y,z) to observation space for each observation to be assimilated and then reflectivity is computed from the interpolated state. Since AI has been applied to each individual variable in the state, the resulting prior reflectivity values incorporate the inflation from all these variables leading to an assimilation system with larger perturbations, increasing the magnitude of AI being applied. This impact is mostly likely to occur in areas of heavy precipitation where perturbations in the hydrometeor variables are greatest.

If the relationship between prognostic and diagnostic variables was completely linear and no AI was applied, then reflectivity (dBZ) interpolated from the state vector (EFO) should be the same as reflectivity computed from the interpolated prognostic variables (IFO). However, since radar reflectivity has a fundamentally nonlinear relationship to prognostic variables, the resulting observational priors differ depending on whether EFO or IFO is employed or whether reflectivity is assimilated in Z or dBZ space. Furthermore, IFO requires the data assimilation system to have a particular microphysics scheme's reflectivity calculation present within the radar forward operator. The current IFO option in DART using reflectivity calculated based on Lin microphysics. For recent GSI-EnKF research, reflectivity is calculated using WSM6 microphysics (Johnson et al. 2015). For this research, both external and internal methods use the Thompson microphysics scheme to calculate reflectivity following development of a Thompson-based reflectivity forward operator for DART (Thompson et al. 2004, 2008).

In order to parallelize the data assimilation scheme, the EAKF system uses the extended state vector method to permit distributed memory parallelism. For both methods, the extended state vector includes the reflectivity priors computed using either EFO or IFO. In the parallel implementation of the EAKF, observations are assimilated serially one after the other. *The interpolated reflectivity priors in observation space are not recalculated from the updated state*, but are directly updated by the data assimilation algorithm as part of the extended state vector. After the new analysis is generated, the posteriors in the observation space are computed via interpolation of the gridded posterior reflectivity field (EFO) or interpolation of the microphysical variables and application of the reflectivity forward operation (IFO). Following this step, observation diagnostics such as innovation, error, and ensemble spread are calculated. Finally, initial conditions for each ensemble member are generated from the filter and used to initiate forecasts for the next assimilation cycle.

4. Idealized model

a. Nature Run

Both the nature run and the idealized data assimilation experiments used by this research are based on those described in Jones and Stensrud (2015). An overview of the previous discussion is provided here. The nature run in which a splitting supercell thunderstorm is simulated using the Weisman-Klemp analytic sounding with a quarter circle hodograph to define the vertical wind profile (Weisman and Klemp 1982, Fig. 2). A stationary horizontal grid using 200 x 150 grid points with a horizontal grid spacing of 1 km is defined for the nature run. The nature run uses 51 vertical levels extending from near the surface to 20 km, with a mean vertical spacing of 400 m. The center of the domain is set to a reference latitude of 35°N and longitude of 97.5°W. A circular thermal warm bubble perturbation potential temperature of 3.0 K, decreasing to 0 K at a horizontal radius of 10 km, is placed at the surface at the center of the domain to initiate convection as in Jones and Stensrud (2015). The nature run is initiated with the

horizontally homogenous atmosphere defined by the input sounding and the warm bubble and is allowed to freely evolve thereafter. The nature run is integrated using version 3.4.1 of the WRF model (Skamarock et al. 2008) over a period of 120 minutes with output saved at 5 min intervals. The nature run utilizes Thompson microphysics (Thompson et al. 2004, 2008), with no radiation or boundary layer schemes applied. Both the nature run and the following experiments use open boundary conditions.
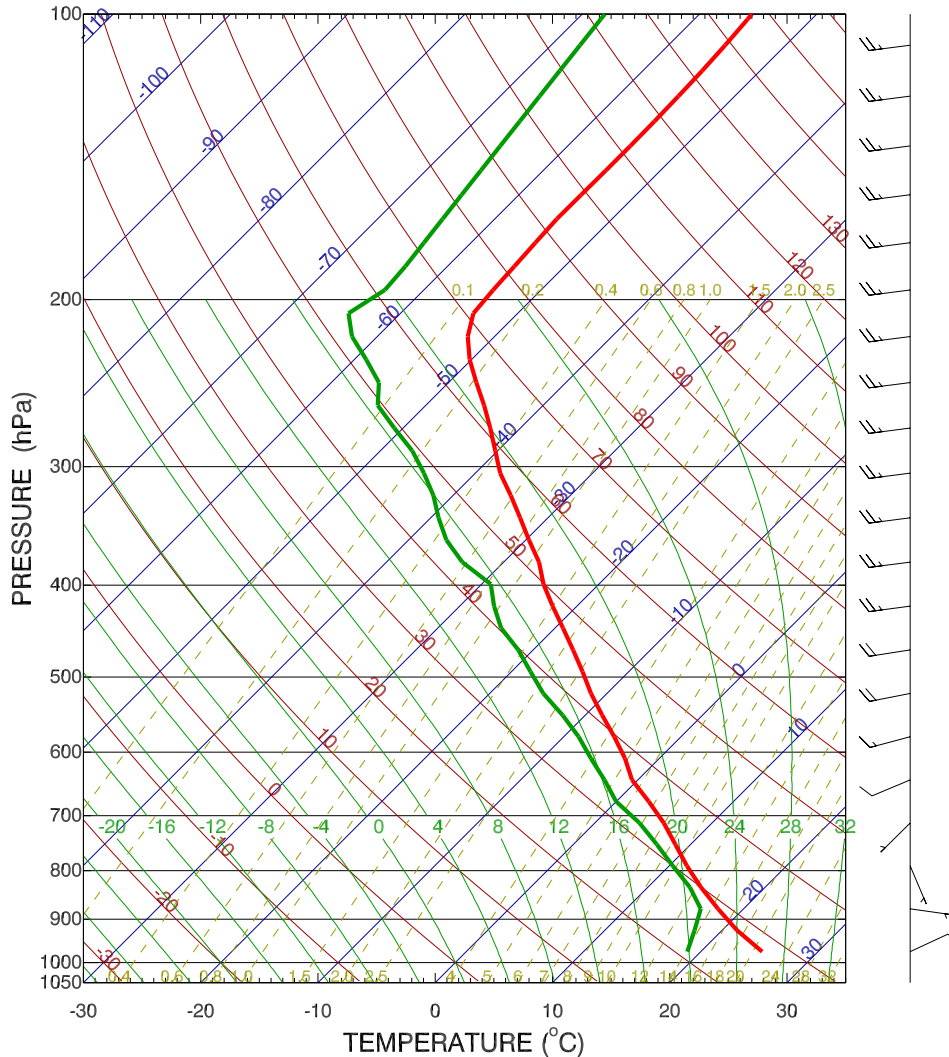

Figure 2 *Weisman – Klemp (1982) sounding profile used to create the environment for the nature run and each experiments. Perturbations are applied to temperature and humidity when each ensemble member is initiated.*

The evolution of the idealized supercell can be shown by tracking the 3 km above ground level (AGL) reflectivity at various times during the nature run (Fig. 3). After only 20 minutes (t=20), a small area of greater than 55 dBZ reflectivity has developed along with a corresponding area of updraft velocity greater than 15 ms^{-1}. By t=40, the area of high reflectivity has grown and the updrafts have split into two cores indicating the onset of a storm split. At t=60 and afterwards, the both storms continue to grow in size and diverge from each other. By t=100, two mature supercells have developed within the nature run, each having separate updrafts and reflectivity cores greater than 70 dBZ.
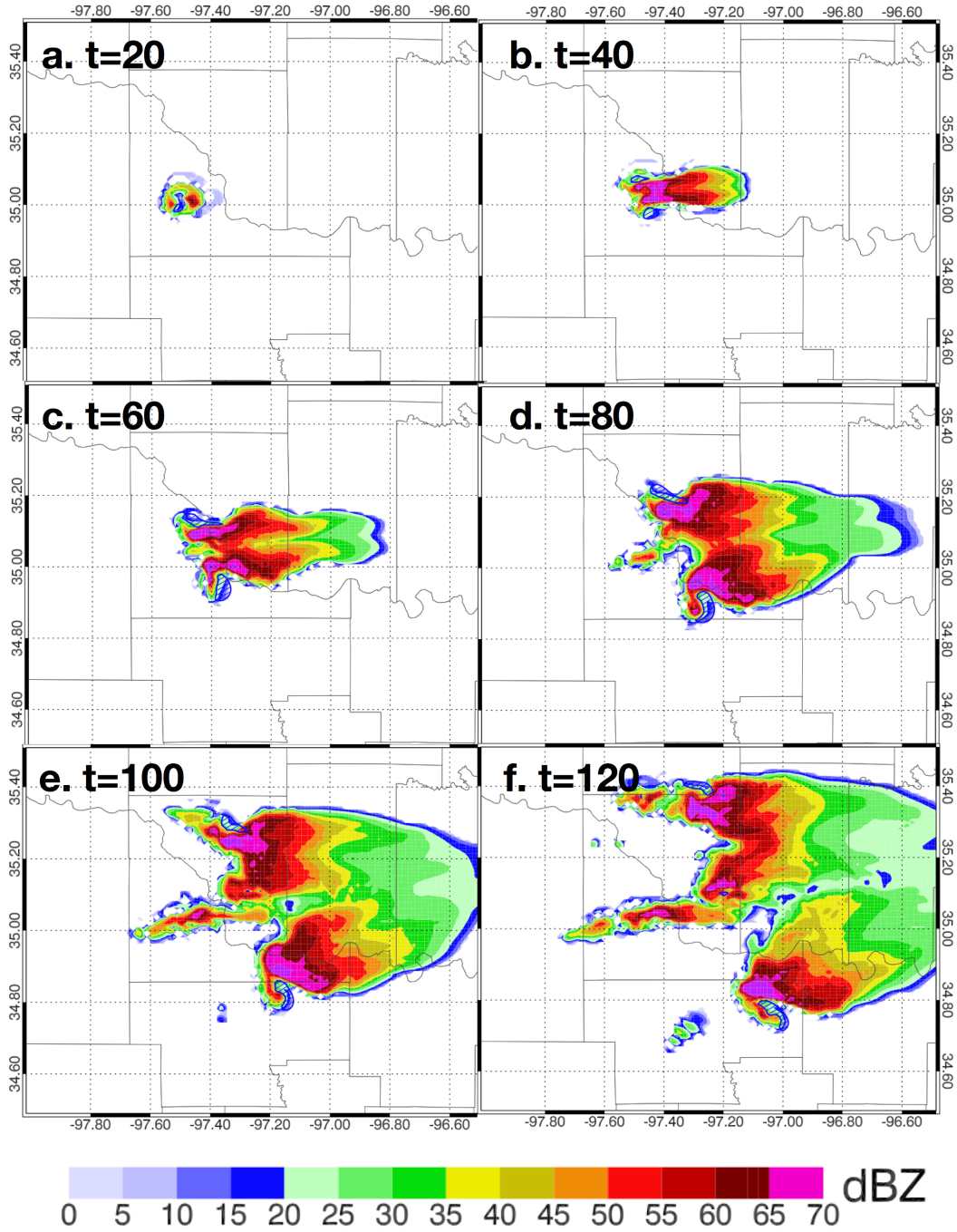
Figure 3. *Nature run 3-km AGL simulated radar reflectivity at 20 (a), 40 (b), 60 (c), 80 (d), 100 (e), and 120 (f) minutes after the warm bubble has been inserted into the WK82 environment. Blue hatched areas indicate areas of 4-km upward vertical velocity greater than 15 ms^{-1}.*

Synthetic radar reflectivity is computed from the nature run by defining a radar location at 97.5°W, 34.3°N and assuming a VCP-21 precipitation mode scan strategy with 9 conical levels (Crum and Alberty 1993). Observation locations are defined using a gate spacing of 4 km with an azimuth resolution of 3 degrees with a maximum observation height of 10 km (Fig. 4). This is significantly coarser than the operational product so that the simulated observations are

larger than the model grid and also to reduce the impact of representativeness error. Reflectivity generated by Thompson microphysics scheme on the model grid is interpolated to the synthetic observation location. Synthetic observations are generated at 5 minute intervals through the duration of the nature run. Observation error for reflectivity is set to 5 dBZ, similar to that used by several previous studies (e.g. Dowell et al. 2004). The selection of the EFO method to create the synthetic radar observations was used since this method was considered the baseline for all the experiments. Calculating simulated observations using the other methods also generated similar results, so the EFO set was selected for all experiments.
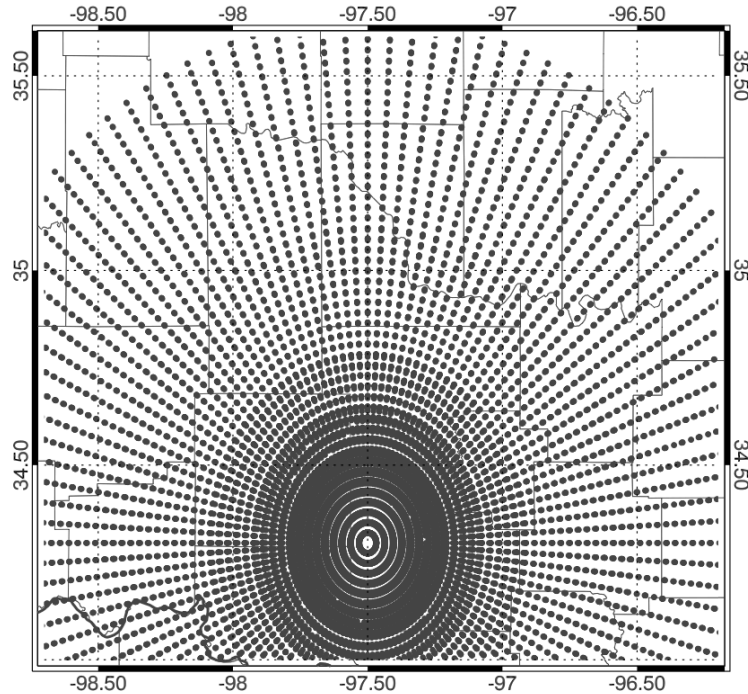

Figure 4*. Plan view of synthetic radar reflectivity observation locations for a simulated radar location at 97.5°W, 34.25°N. Note that observation resolution decreases as distance from the radar increases.*

b. Data assimilation

The EAKF approach assimilates synthetic radar reflectivity observations into a horizontally homogeneous initial state similar to that used by the nature run. Temperature and moisture perturbations are applied to the previous input sounding in order to generate initial conditions for a set of 40-ensemble members as in Jones and Stensrud (2015). The random perturbations have a Gaussian distribution with a standard deviation of 1 K for potential temperature and 0.5 g kg^{-1} for water vapor mixing ratio centered on the original sounding value and out to a maximum of two standard deviations at the surface and decreasing to near zero at the top of the sounding (20 km). For water vapor mixing ratio, any negative values are set to 0.01 g kg^{-1} and any values causing super saturation are reduced to a value equal to saturation. No warm bubble is inserted in the initial conditions. As such, the development of convection within each experiment requires the assimilation of the synthetic radar reflectivity observations to initiate storms.

The EAKF present in the DART software (Anderson and Collins 2007; Anderson et al. 2009) coupled with WRF version 3.4.1 is used to assimilate observations and integrate forecasts

for the 40 member ensemble. The characteristics of the data assimilation system are very similar to those described in Jones and Stensrud (2015). The horizontal domain is centered at the same location as the nature run, but now has a larger 2 km horizontal grid spacing encompassing 150 x 110 grid points. The vertical grid spacing and number of vertical levels are the same as those used in the nature run. Synthetic radar reflectivity observations greater than 10 dBZ are assimilated at 10 minute intervals for a period of 60 minutes at which time 60 minute forecasts are initiated for each member. For these observations, horizontal and vertical localization radii of 18 km and 6 km respectively are applied using the Gaspari and Cohn (1999) technique. Adaptive inflation is applied to the prior state of each assimilation cycle to maintain appropriate spread in the model and prevent filter divergence (Anderson 2007, 2009). Initial prior inflation factor (inf_initial) is set to 1.0 with a standard deviation (inf_sd_initial) of 0.6. No outlier thresholds are applied to ensure the same number of observations are assimilated in all experiments.

To test the impact of interval vs. external radar reflectivity assimilation and whether or not AI magnifies this impact, 6 experiments are conducted using this data assimilation system (Table 2). The first uses the EFO method with no adaptive inflation applied whereby reflectivity calculated from the Thompson code inside WRF 3.4.1 and included in the WRF output files is interpolated to the observation location to create the priors (EFO-N). A similar experiment is conducted in which reflectivity is interpolated in *Z* space instead of dBZ space (EFO-NZ). Adaptive inflation is then applied to both forming EFO-A and EFO-AZ respectively. The final two experiments use the internal method in which the Thompson microphysics code is added to the forward operator to calculate reflectivity from the interpolated state with (IFO-A) and without (IFO-N) adaptive inflation applied. By removing the adaptive inflation from one set of experiments, it will be possible to isolate the differences in the analysis due to just the interpolation from those due to the combined effects of interpolation and adaptive inflation.

Experiment	Forward Op.	Reflectivity	AI
EFO-N	External	dBZ	NO
EFO-NZ	External	Z	NO
EFO-A	External	dBZ	YES
EFO-AZ	External	Z	YES
IFO-N	Internal	dBZ	NO
IFO-A	Internal	dBZ	YES

Table 2. List of experiments performed for both idealized and real-case (20 May 2013) data sets.

5. Idealized experiments

a. Reflectivity and hydrometeor characteristics

The number of assimilated observations increases from less than 600 at the first assimilation cycle at t=10 to 1200 at the final assimilation cycle at t=60 as the storm matures in the nature run (not shown). By t=60, all experiments have generated strong convection in the model as shown by the 3 km AGL model simulated radar reflectivity analysis at this time (Fig. 5). All experiments are broadly similar in their characterization of reflectivity, but several small differences are apparent. EFO-AZ is the only experiment that fails to generate reflectivity > 65 dBZ at this level and produces the lowest spatial coverage of reflectivity > 55 dBZ (Fig. 5d). Conversely, the IFO experiments generate larger areas of reflectivity > 55 dBZ than any of the EFO experiments with IFO-A generating the highest reflectivity values of all experiments (Fig. 5e, f).
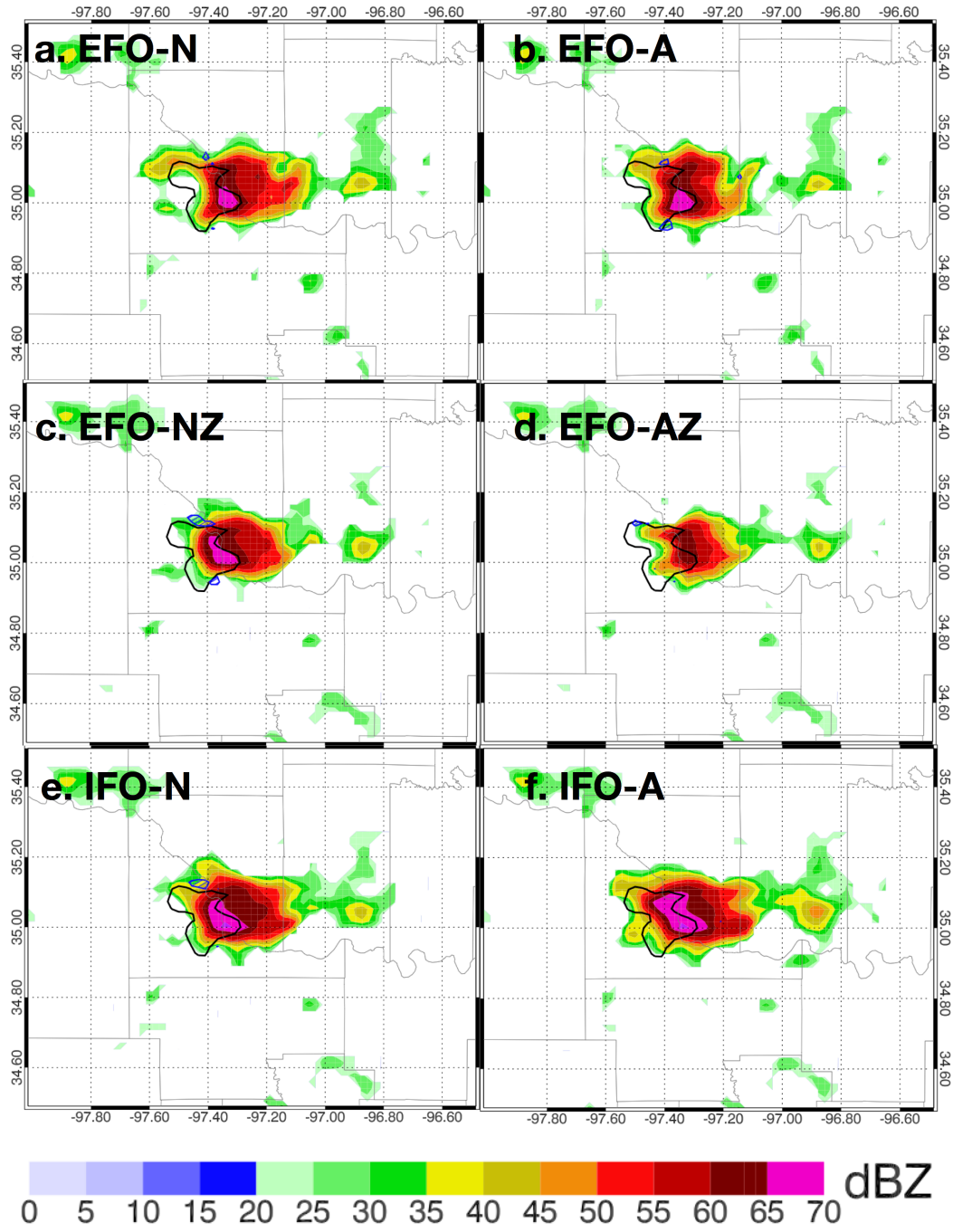
Figure 5. *Ensemble mean 3-km AGL simulated reflectivity for each experiment at t=60. Blue hatched area indicates location where 4 km upward vertical velocity exceeds 15 ms^{-1}.*

The differences in the reflectivity characteristics of these experiments can be explained by analyzing the distributions of selected hydrometeor mixing ratio concentrations. Figure 6a shows the number of ensemble mean rain water mixing ratio (QRAIN) grid points at all model levels within 0.25 g kg^{-1} bins for values ranging from 1.0 to 8.0 g kg^{-1} at t=60. For all experiments, the number of grid points within each bin decreases rapidly as QRAIN increases. Differences between the non-AI experiments are small though IFO-N generates slightly more > 3

g kg^{-1} grid points than EFO-N while EFO-NZ generates the least. Once AI is applied, the differences become more apparent. EFO-A is still similar to the non-AI experiments, but now EFO-AZ generates far fewer > 2 g kg^{-1} QRAIN grid points compared to the other experiments. Conversely, IFO-A generates larger numbers of > 4 g kg^{-1} grid points than any of the other experiments. Figure 6b shows a similar distribution for graupel mixing ratio (QGRAUP) at t=60. As with QRAIN, the non-AI experiments are similar though EFO-NZ does generate somewhat less QGRAUP than EFO-N or IFO-N. Once AI is applied, the differences are magnified significantly. EFO-AZ generates the least QGRAUP while IFO-A generates the most, with the magnitude of the differences greater than observed for QRAIN. Differences in the concentrations for other hydrometeor species are generally smaller (not shown).
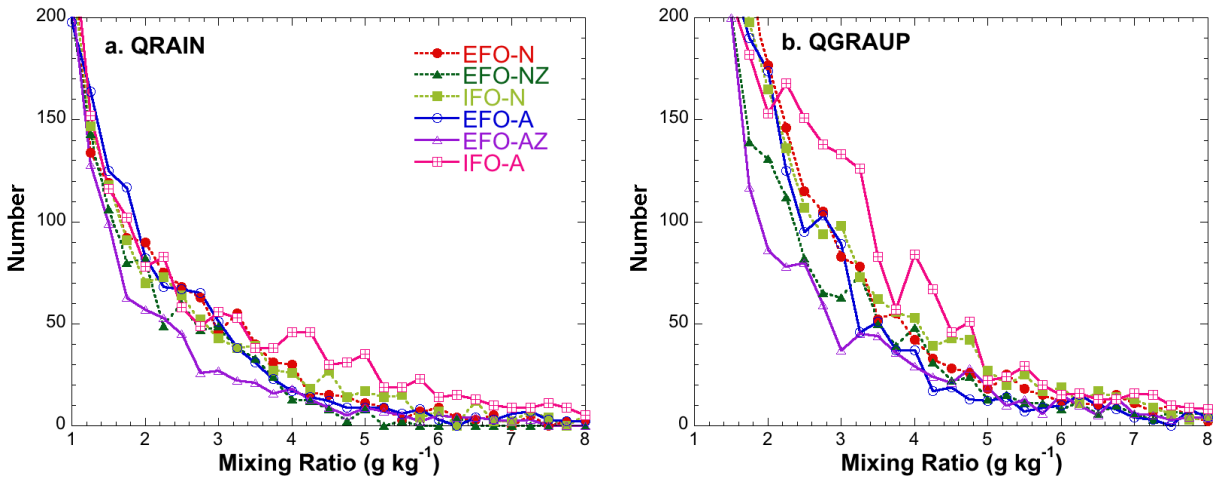

Figure 6. *Frequency distributions of ensemble mean QRAIN (a) and QGRAUP (b) mixing ratios at t=60 for each experiment. Bin size is 0.25 g kg^{-1}.*

These differences can also be visualized by analyzing the vertical distribution of hydrometeor mixing ratios associated with the analyzed convection. The average QRAIN and QGRAUP from the ensemble mean analysis of each experiment are calculated for all grid points where reflectivity > 25 dBZ at each model level (Fig. 7a,b). Figure 7c provides the corresponding number of grid points > 25 dBZ used to calculate the averages for each experiment. Unlike the histogram distributions described above, the vertical profiles of QRAIN are noticeably different in the non-AI experiments (Fig. 7a). Below 800 hPa, EFO-N generates ~1.1 g kg^{-1} at each model level. EFO-NZ generates less, 1.0 g kg^{-1}, while IFO-N generates more, 1.25 g kg^{-1}. Between, 800 – 600 hPa, EFO-N and IFO-N becomes similar while EFO-NZ consistently generates less QRAIN. Differences between methods are magnified once AI is applied with IFO-A generating > 1.7 g kg^{-1} in QRAIN at each model level below 800 hPa (Fig. 7a). Both EFO-A and EFO-AZ slightly decrease QRAIN in this layer compared to the corresponding non-AI experiments. Above 800 hPa, applying AI to the external method does increase QRAIN compared to the corresponding non-AI experiments, while IFO-A continues generating the largest QRAIN concentrations.

Additional differences are apparent when comparing the vertical profiles of QGRAUP (Fig. 7b). QGRAUP generally increases with height up to near 500 hPa. Between 500 hPa and the surface, all experiments generate a similar QGRAUP profile with EFO-A and EFO-NZ generating the lowest concentrations and IFO-N and IFO-A generating the highest. Above 500

hPa larger differences between each experiment begin to emerge. In the 500 – 300 hPa layer, IFO-N generates a slightly higher average QGRAUP than EFO-N, which itself is higher than EFO-NZ. Applying AI again increases the differences between methods with IFO-A generating the largest average QGRAUP with EFO-AZ generating the least (Fig. 7b). Comparing the number of grid-points > 25 dBZ as a function of height also shows large differences between the experiments (Fig. 7c). IFO-A generates the highest number with over 300 grid points per level above 800 hPa. EFO-N is similar followed by IFO-N, EFO-A, EFO-NZ, and EFO-AZ. The latter two experiments also consistently generated the lowest average QRAIN and QGRAUP indicating that interpolating reflectivity factor instead of dBZ results smaller, weaker storms in the analysis. The high number of grid points for EFO-N was somewhat surprising, but further investigation revealed that this experiment generated larger areas of 25-35 dBZ reflectivity indicating larger areas of non-convective precipitation in this experiment. The corresponding QRAIN, QGRAUP, and number of grid-points > 25 dBZ from the nature run at t=60 are also provided for reference. In general, the nature run has lower QRAIN concentrations than most of the experiments while having similar to higher average QGRAUP compared to the experiments.
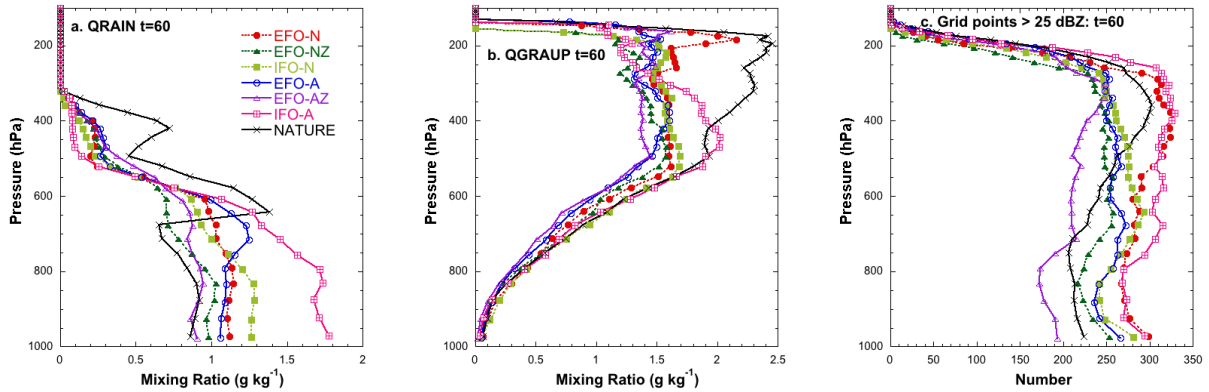

Figure 7. *Vertical profiles of ensemble mean QR (a) and QG (b) averaged at each model level over all grid points where analyzed reflectivity > 25 dBZ at t=60. Number of grid points > 25 dBZ as a function of height is provided in panel c. Dots indicate value at each model level.*

While these results have shown that small differences exist due to changes in the reflectivity forward operator alone, it is clear that applying AI magnifies these differences. To better assess this the influence of AI, a frequency distribution of the prior AI factor applied to QRAIN and QGRAUP at t=60 is calculated (Fig. 8). The highest values correspond to the areas of highest QRAIN and QGRAUP in the prior mean analysis. Several differences exist in the AI distributions of QRAIN between EFO-A, EFO-AZ, and IFO-A experiments (Fig. 8a). For AI values between 3 and 4, EFO-A generates the greatest number of grid points while EFO-AZ generates the least. Between 4 and 7, all experiments are similar. For values higher than 7, IFO-A consistently has the greatest number. Similar patterns exist for QGRAUP where EFO-A generates the highest numbers between 3 and 5, with IFO-A generating higher number of grid points where AI > 7. The goal of AI is to increase spread within the model, and the distribution and magnitude of this spread will impact the resulting model analysis. Since IFO-A generates the highest prior AI, the increase in spread in the strongest portion of the analyzed convection is also the greatest. The greater spread allows larger convariances between reflectivity and hydrometeor variables to exist, increasing their impact when assimilated. Conversely, EFO-AZ generates the

15

smallest prior AI factors resulting in less spread and a lower impact from reflectivity assimilation.
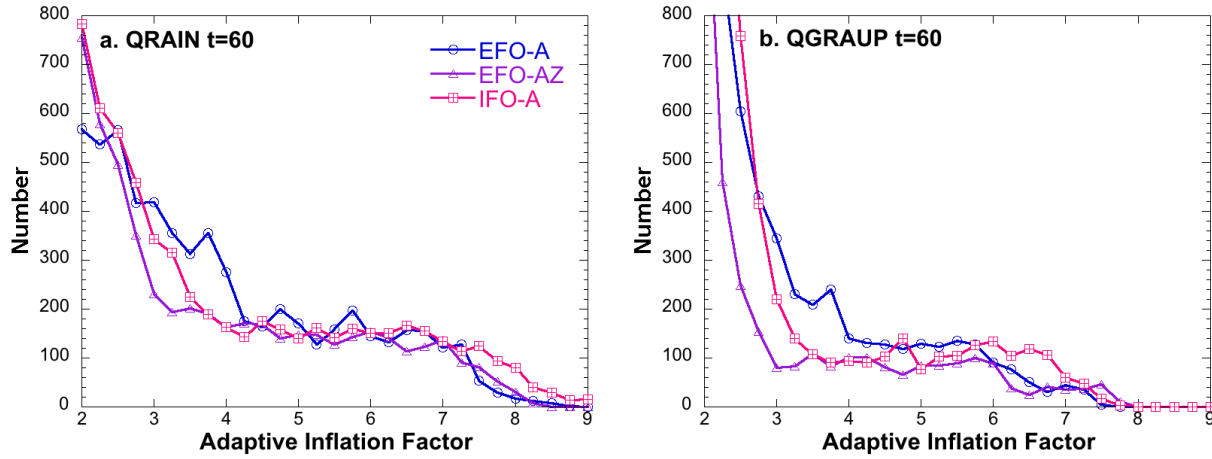

Figure 8. *Frequency distributions of prior adaptive inflation factor at t=60 for QRAIN (a) and QGRAUP (b). Bin size is 0.25.*

b. Cold pool and precipitation characteristics

The importance of adaptive inflation and reflectivity forward operator method are also evident when comparing the cold-pool characteristics from each experiment. At t=60, all experiments have generated a cold pool associated with the analyzed convection (Fig. 9). For non-AI experiments, IFO-N generates the coldest surface temperatures followed closely by EFO-N and finally EFO-NZ. Applying AI does not further cool the the cold pools, but does alter their shape and spatial coverage. In general, these experiments are somewhat more "noisy" with the cold pool becoming intermixed with areas of warmer surface temperatures causing large temperature gradients in the analysis. This is most evident in the EFO-A and EFO-AZ experiments, but is also noticeable in IFO-A (Fig. 9b,d,f).
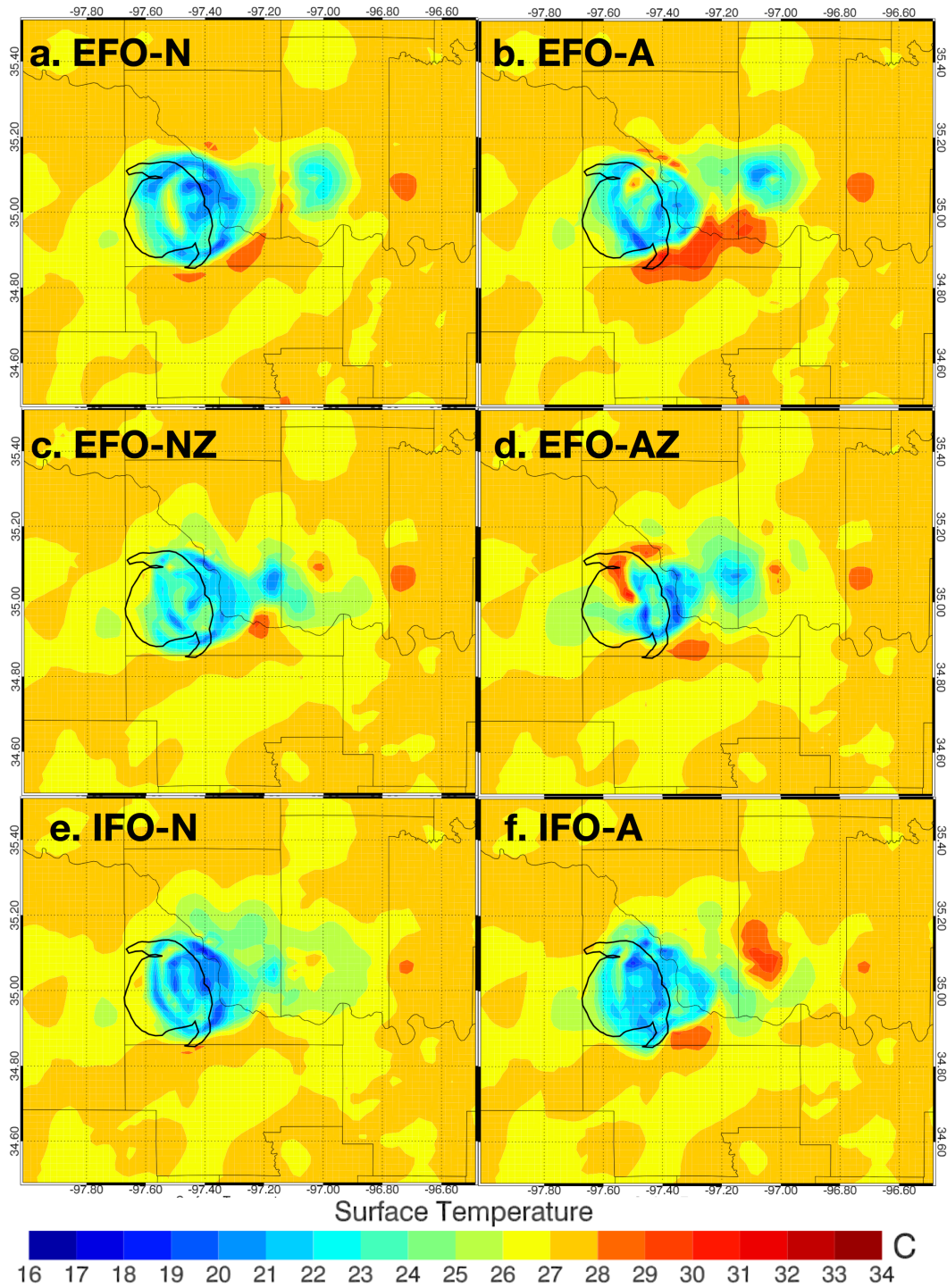
Figure 9. *Ensemble mean surface temperature for each experiment at t=60. Nature run temperature less than 24°C is denoted by the black contour.*

Differences in the analyses between each experiment have a significant impact on the ensuing forecast. To determine the immediate impact, the ensemble mean accumulated precipitation for a 10 minute forecast beginning at t=60 is shown in Figure 10. All experiments generate precipitation corresponding to the nature run, but have a general eastward bias. The EFO-NZ and EFO-AZ experiments generate the least amount of precipitation during this period,

consistent with the hydrometeor analysis above. Similarly, IFO-A generates the greatest amount of precipitation and is also the only experiment where the forecast precipitation maxima match well with those in the nature run (Fig. 10f).
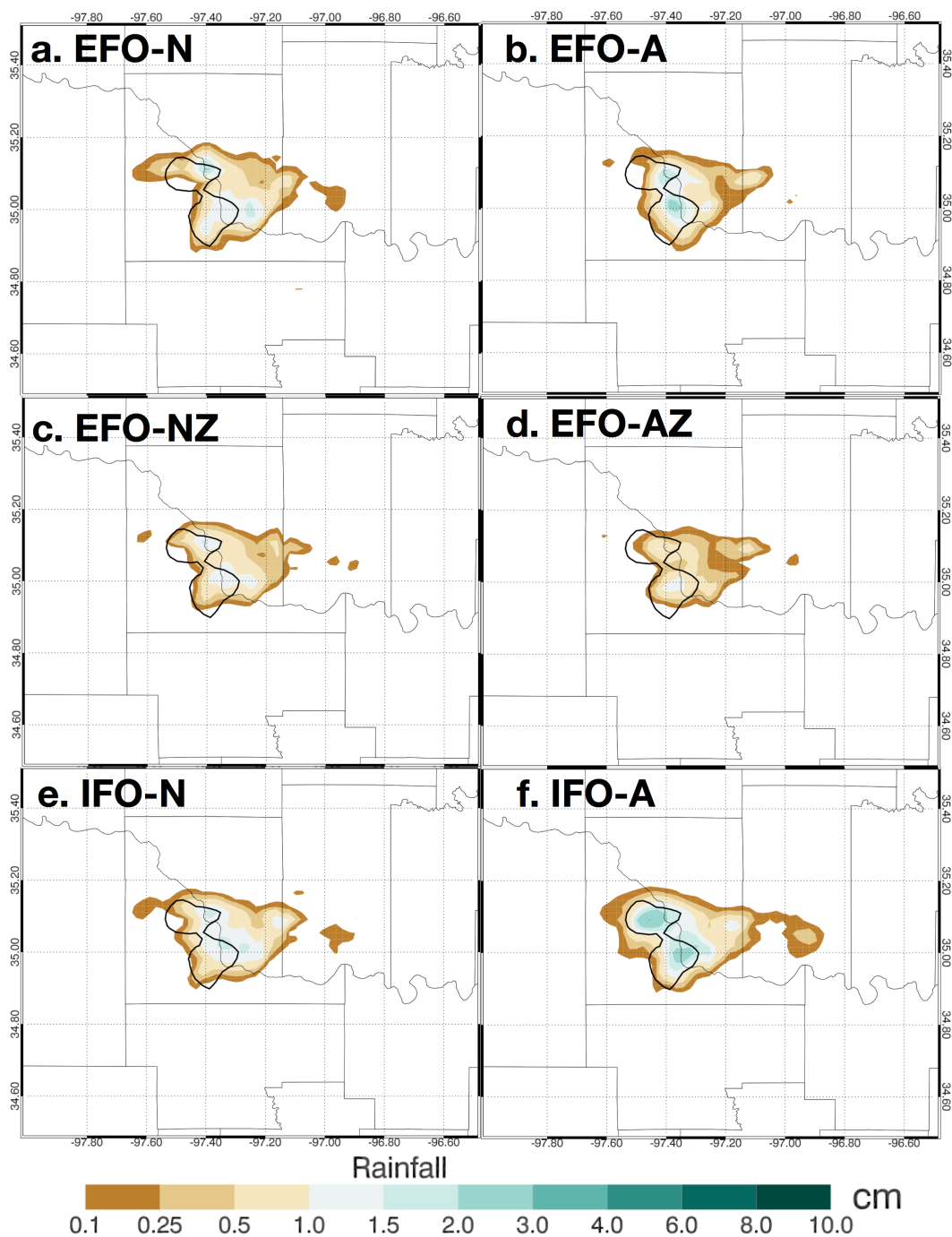

Figure 10. *Ensemble mean 10 minute accumulated precipitation between t=60 and t=70. The nature run precipitation greater than 2 cm is denoted by the black contour.*

To quantify the differences in precipitation from each experiment relative to the nature run, the Equitable Threat Score (ETS) (Wilks 2006) is computed for accumulated precipitation

beginning at t=60 and continuing out to t=120. For each ensemble member, if the experiment generates accumulated precipitation greater than a threshold value at a forecast time within ±6 km of the nature run accumulated precipitation ending at the same time, this is considered a "hit". If the experiment generates accumulated precipitation greater than the threshold and this threshold is not exceeded in the nature run, then it is considered a false detection. Finally, if neither the nature run nor the experiments exceed this threshold, then it is considered a correct null forecast. The "mean ETS" is calculated by averaging ETS over all ensemble members for a particular forecast time. ETS is calculated for 5 thresholds (0.5, 0.75, 1.0, 1.5, and 2.0 cm) to assess the difference in model skill in light vs. heavy precipitation (Fig. 11).
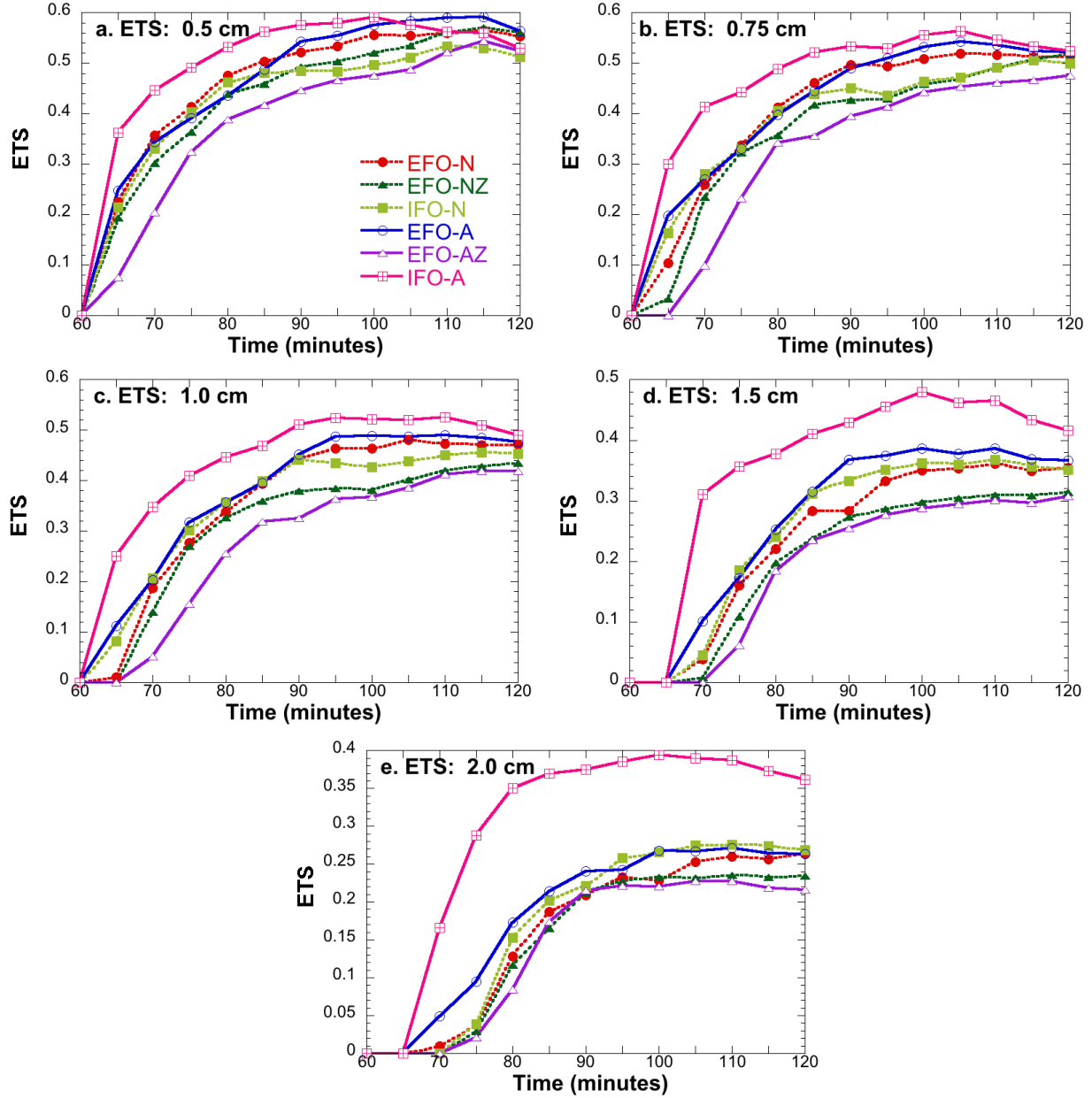

Figure 11. *ETS for accumulated precipitation between the nature run and forecasts ranging from 5 minutes to 1 hour (t=65 min to t=120 min after the start of the nature run) for various precipitation thresholds.*

For all experiments, ETS increases as a function of time as the accumulation period increases. The highest ETS (~0.6) occurs using the 0.5 cm threshold with maximum values decreasing as the precipitation threshold increases. ETS for non-AI experiments are generally similar with EFO-N having somewhat higher skill for thresholds ≤ 1 cm and IFO-N having higher skill for the higher thresholds. This indicates that EFO-N does a good job of forecasting the light to moderate precipitation, but is less likely to generate heavy precipitation. EFO-NZ generally has the least skill of the three non-AI experiments. As before, the application of AI increases the differences between various reflectivity assimilation methods. For forecasts up to 30 minutes, IFO-A generates the highest ETS compared to the other experiment for the 0.5 cm precipitation threshold. For higher thresholds, the increase in skill from IFO-A exists out to at least 1 hour though decreasing sample size reduces the statistical significance of this increase. EFO-A is similar to EFO-N for the lower precipitation thresholds, but EFO-A does increase skill for higher thresholds. Conversely, applying AI in EFO-AZ lowers skill relative to EFO-NZ and consistently has the lowest skill of the six experiments conducted.

The differences between hydrometeor, temperature, and precipitation characteristics between each experiment during the assimilation period indicate that the selection of reflectivity forward operator method and adaptive inflation is very important. When no adaptive inflation is applied, small differences in the characteristics of convection exist between the external and internal reflectivity assimilation methods. Once adaptive inflation is applied, the magnitude of these differences increases significantly. These results suggest that adaptive inflation is magnifying the inherent non-linearity present in the calculation of reflectivity. Further analysis using larger samples sizes and real data sets is necessary to determine if these differences carry over into the real world.

6. 20 May 2013 Event

a. Case study and experiment characteristics

Atmospheric conditions were highly favorable for tornadic supercells on 20 May 2013 in central Oklahoma. Convection was rapidly developing along the dryline by 1900 UTC and became tornadic within the next hour. To analyze this event, this research uses the ensemble data assimilation system similar to the one described in detail by Wheatley et al. (2015) and Jones et al. (2016) known as the NSSL experimental WoF System for ensembles (NEWS-e). The NEWS-e uses WRF-ARW version 3.6.1 (Skamarock et al. 2008) coupled with the EAKF in DART (Anderson and Collins 2007; Anderson et al. 2009) to assimilate radar observations at 15 minute intervals in a 36-member ensemble.

The experimental domain utilizes a 1-way nest setup, whereby the parent and nested grids are run concurrently with information only exchanged at the lateral boundaries at 1-hour intervals once storm-scale data assimilation commences (with no feedback from the nested grid to the parent grid). The parent domain has a horizontal grid spacing of 15-km covering the continental United States, while the nested domain has a horizontal grid spacing of 3-km centered in central Oklahoma. Both parent and nested domains have 56 vertical levels ranging from the surface to a model top of 10 hPa. Initial conditions for the parent and nested grids are downscaled from the 21-member 0000 UTC 19 May Global Ensemble Forecast System (GEFS) forecast cycle. The GEFS also provides boundary conditions for the parent grid. Different sets of WRF model physics options are applied to each ensemble member to account for model physics

uncertainties (e.g. Stensrud et al. 2000). All members use Thompson cloud microphysics (Thompson et al. 2004, 2008) and no cumulus parameterization is applied on the storm-scale grid (Jones et al 2013; Jones et al. 2015; Wheatley et al. 2015).

The mesoscale ensemble is integrated hourly, assimilating conventional observations including Mesonet, METAR, marine, radiosonde, ACARS, and satellite winds as described in Wheatley et al. (2015). Horizontal and vertical localization radii of ~458 km and 6 km respectively are applied using the Gaspari and Cohn (1999) technique. Available Mesonet observations are only assimilated in the storm-scale domain using a smaller horizontal localization radius of 60 km. For this event, 15 minute cycling of radar data assimilation on the 3 km nested grid begins at 1800 UTC and continues through to 2000 UTC, after which 1 hour ensemble forecasts are generated. Level 2 radar reflectivity and radial velocity observations are obtained from 3 Oklahoma radars (KFDR, KTLX, KINX) with observations reanalyzed to a 6 km horizontal resolution grid using a Cresman analysis technique (Majcen et al. 2008). The number of radar reflectivity observations, not including clear-air reflectivity, assimilated increases from < 1000 at 1800 UTC to > 10000 at 2000 UTC. The number of clear-air reflectivity observations assimilated ranges between 12000 and 26000 (not shown).

Horizontal and vertical localization radii of 18 and 6 km are applied to both radar reflectivity and radial velocity observations. As in the idealized case, 6 experiments are conducted using the real data case. Experiments include EFO-N, EFO-NZ, and IFO-N, which represent the external method that interpolates dBZ, the external method that interpolates Z, and the internal method where reflectivity is calculated from the interpolated state each with no AI applied. The second set (EFO-A, EFO-AZ, and IFO-A) are similar, except that AI is now applied to each. Settings for AI are the same as those used for the idealized experiments. Additive noise is applied in all experiments to perturb the state where reflectivity > 25 dBZ in the fashion described by Wheatley et al. (2015).

Unlike the idealized case, Doppler radial velocity is also being assimilated in the real data experiments. Previous research (Gao et al. 1999; Gao and Stensrud 2012) used a fall velocity adjustment to account for the fall speed of hydrometeors in the calculation of radial velocity in the forward operator. The adjustment is maximized when comparing against radial velocity observations from higher scan angles. It has been found that including this adjustment in the state vector or calculating it within the forward operator can also significantly impact the analysis, especially near radar sites where high elevation scans of convection are common. In order to isolate the differences due to reflectivity alone, the 20 May experiments use the same fall velocity calculation inside the forward operator. The DART software includes a function based on interpolated density and reflectivity at the observation location to estimate fall velocity. Thus, the EFO calculates fall velocity from the interpolated reflectivity from the state vector and IFO calculates fall velocity from the reflectivity calculated inside the forward operator. (Other combinations are possible including using a state vector fall velocity in both or calculating the fall velocity from the Thompson microphysics code added to the reflectivity forward operator).

b. *Reflectivity and hydrometeor characteristics*

Following 9 assimilation cycles ending at 2000 UTC, all experiments generate several supercells in southern and central Oklahoma as shown by the ensemble mean of model simulated radar reflectivity at 3 km AGL (Fig. 12). Observed reflectivity from the Multi-radar, Multi-sensor (MRMS) product > 45 dBZ is overlaid to compare against model analyses. Several differences are apparent between the experiments. EFO-N fails to generate reflectivity > 50 dBZ

corresponding with the northern storm (#1) which went on to produce the tornado in Moore, Oklahoma (Fig. 12a). Much higher values of reflectivity were generated for the southern OK storms (#2, #3). EFO-A is similar with the southern storms, but generates a much better representation of #1 than does EFO-N. EFO-NZ and EFO-AZ are generally similar with both generating reflectivity cores > 60 dBZ for all storms; however, the coverage of high reflectivity values is somewhat smaller in EFO-AZ, which is consistent with the EFO-NZ and EFO-AZ idealized experiments (Fig. 5,12). Finally, the IFO experiments generate the highest reflectivity values associated with #1 with IFO-A generating the highest values of any experiment for all storms (Fig. 12e,f).
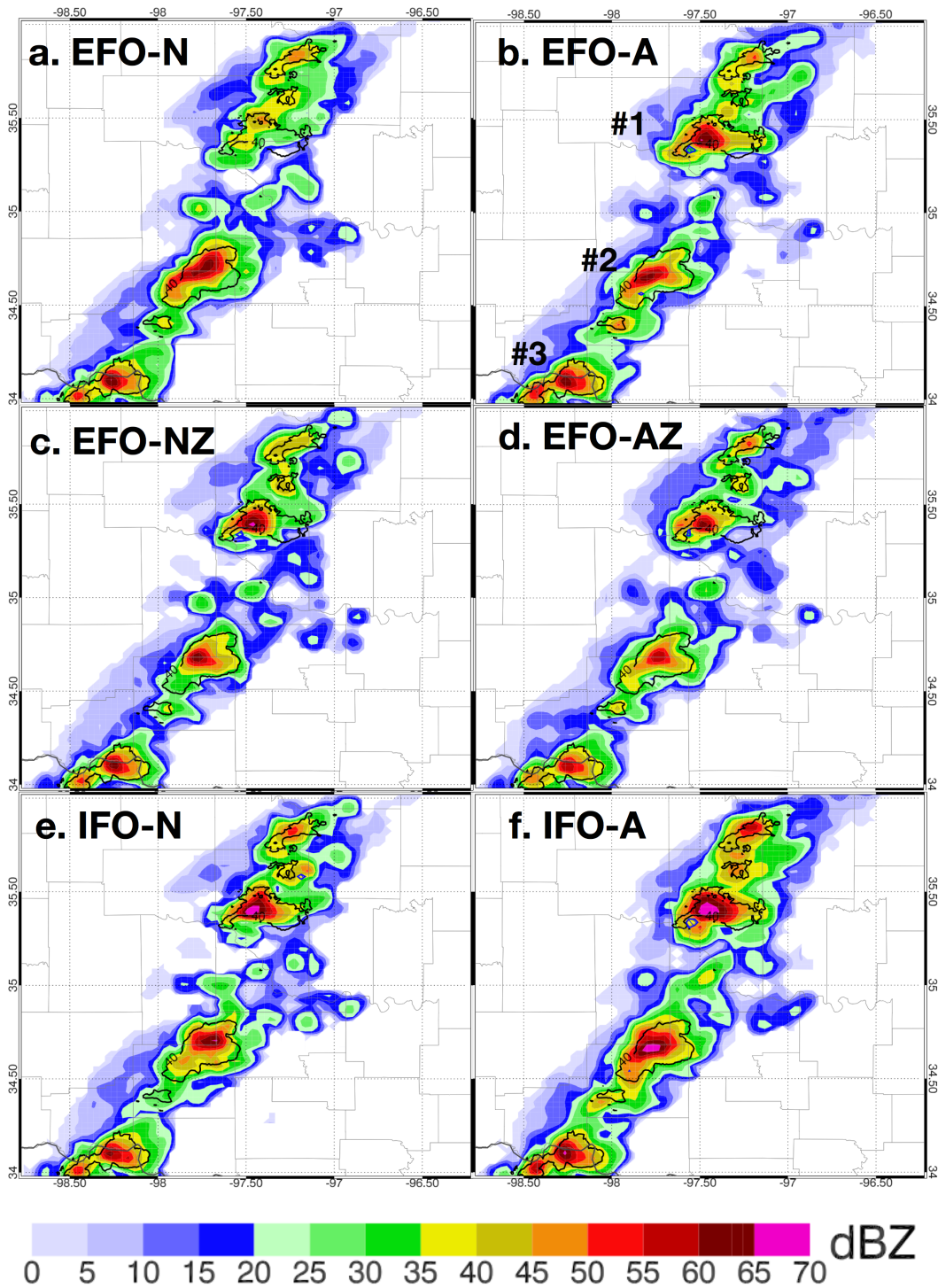
Figure 12. *Ensemble mean 3-km AGL reflectivity at 2000 UTC 20 May 2013 for each experiment.*

To determine if these differences carry over to the hydrometeor characteristics of the analyzed convection, distributions of QRAIN and QGRAUP for the ensemble mean analysis at 2000 UTC are computed in the same manner as in the idealized experiments (Fig. 13a,b). Both EFO-N and EFO-NZ generate similar QRAIN distributions, which are both less than IFO-N indicating that differences between the internal and external methods alone are having a measureable impact on the model hydrometeor analysis (Fig. 13a). As before, applying AI increases the differences between each assimilation method, with EFO-AZ generating the fewest number of QRAIN > 1.0 g kg^{-1} grid points and IFO-A generating by far the most. QGRAUP is similar except that the differences between the non-adaptive inflation experiments are smaller (Fig. 13b).
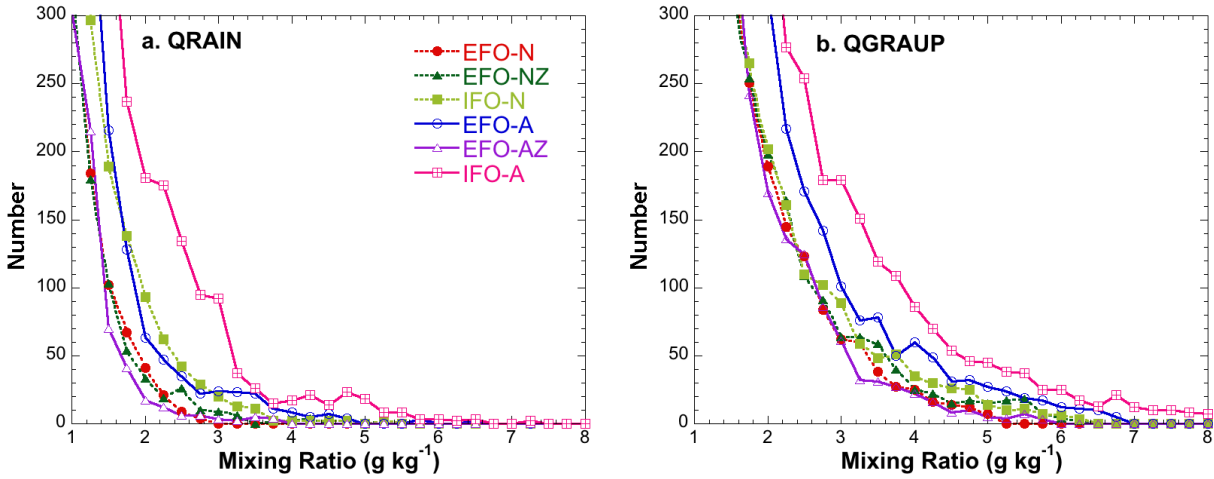

Figure 13. *Frequency distributions of ensemble mean QRAIN (a) and QGRAUP (b) mixing ratios at 2000 UTC 20 May for each experiment. Bin size is 0.25 g kg^{-1}.*

To assess these differences as a function of height, the average QRAIN and QGRAUP are calculated at each model level for grid-points where analyzed reflectivity > 25 dBZ for each experiment at 2000 UTC (Fig. 14). All experiments generate a maximum in QRAIN near 750 hPa with values decreasing slightly towards the surface (Fig. 14a). QRAIN becomes small above 500 hPa as most hydrometeors at these levels are of the frozen variety. A clear separation exists between the non-AI experiments with EFO-N generating the lowest values and IFO-N generating the highest, with an average increase of ~0.1 g kg^{-1} over EFO-N and 0.05 g kg^{-1} over EFO-NZ. Applying AI enhances these differences, but in two different ways. EFO-A and IFO-A generate higher QRAIN concentrations then their non-AI counterparts. However, EFO-AZ generates less QRAIN than EFO-NZ. Similar results were noted in the idealized experiments where applying AI to the experiments that interpolates radar reflectivity factor also resulted in a decrease in analyzed hydrometeor concentrations.
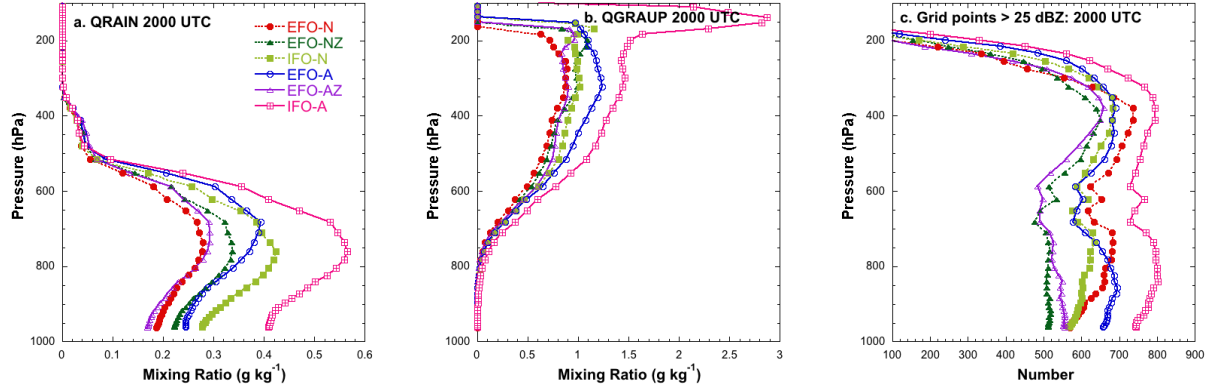
Figure 14. *Vertical profiles of ensemble mean QR (a) and QG (b) averaged at each model level over all grid points where analyzed reflectivity > 25 dBZ at 2000 UTC 20 May. Number of grid points > 25 dBZ as a function of height is provided in panel c. Dots indicate value at each model level.*

For all experiments, QGRAUP gradually increases as a function of height staring from near zero at 800 hPa to values in excess of 1.0 g kg^{-1} above 400 hPa (Fig. 14b). Differences between the non-AI experiments are smaller, but IFO-N does generate somewhat higher values in the 600 – 200 hPa layer. As before, applying AI often magnifies the differences between experiments with EFO-A and especially IFO-A generating much higher QGRAUP concentrations than EFO-AZ. The number of grid points > 25 dBZ also differs between experiments with IFO-A having the most and EFO-NZ and EFO-AZ having the least (Fig. 14c). The much smaller number of precipitation-generating grid-points in EFO-NZ, compared to EFO-N or IFO-N, could be a factor in why applying AI actually reduces hydrometeor concentrations rather then increases them.

To assess the impact of AI, distributions of prior adaptive inflation factor are plotted for QRAIN and QGRAUP (Fig. 15). For both QRAIN and QGRAUP, IFO-A has the greatest number of grid points where AI > 2.5. EFO-A and EFO-AZ are closer, but differ in a couple of important aspects. First, EFO-A has the least number of AI grid points < 3.25 for both QRAIN and QGRAUP. However, EFO-A generates ~50 AI > 4 grid points whereas EFO-AZ generates virtually none. This corresponds to the overall smaller number of precipitating grid-points in this experiment and suggests that a prior analysis with fewer number of precipitating grid-points results in less prior AI applied to the state, which reduces spread in these regions. The reduction in spread reduces covariances between observations and the model state, limiting their impact and inhibiting optimal assimilation of the data. Conversely, the larger values from IFO allow for higher covariances and more optimal assimilation, generating stronger storms over multiple assimilation cycles.
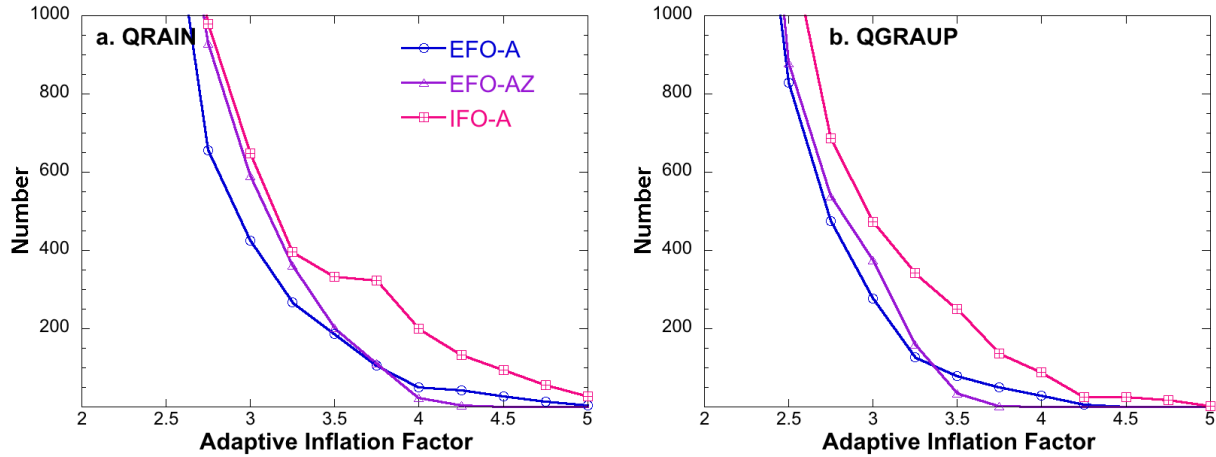
Figure 15. *Frequency distributions of prior adaptive inflation factor at 2000 UTC 20 May for QRAIN (a) and QGRAUP (b). Bin size is 0.25.*

c. Cold pool analysis

The differences in hydrometeor characteristics between each experiment also impact the cold-pool characteristics of the convection in the model analysis. Figure 16 shows ensemble mean surface temperature at 2000 UTC for each experiment. The magnitude of the cold pools differs between experiments with EFO-N having the weakest (warmest) and IFO-A having the strongest (coldest). This is consistent with the differences in hydrometeor concentrations observed between each experiment with larger QRAIN concentrations corresponding to colder surface temperatures. Limited verification of the surface temperature analysis is possible by comparing against OK Mesonet observations. Table 3 shows root mean square error (RMSE) between ensemble mean surface temperature at 2000 UTC and the Mesonet observations at 33 sites within the plotted domain. The non-AI experiments generate an RMSE ~2.0 °C while the AI experiments generally have lower errors. IFO-A generates the lowest RMSE at 1.7 °C with the next lowest experiment being EFO-A. The improvement primarily results from the colder temperatures generated around the edges of the storms in the AI experiments. Unfortunately, there were no observations near the center of the cold pools to validate which experiment correctly analyzed their maximum strength.
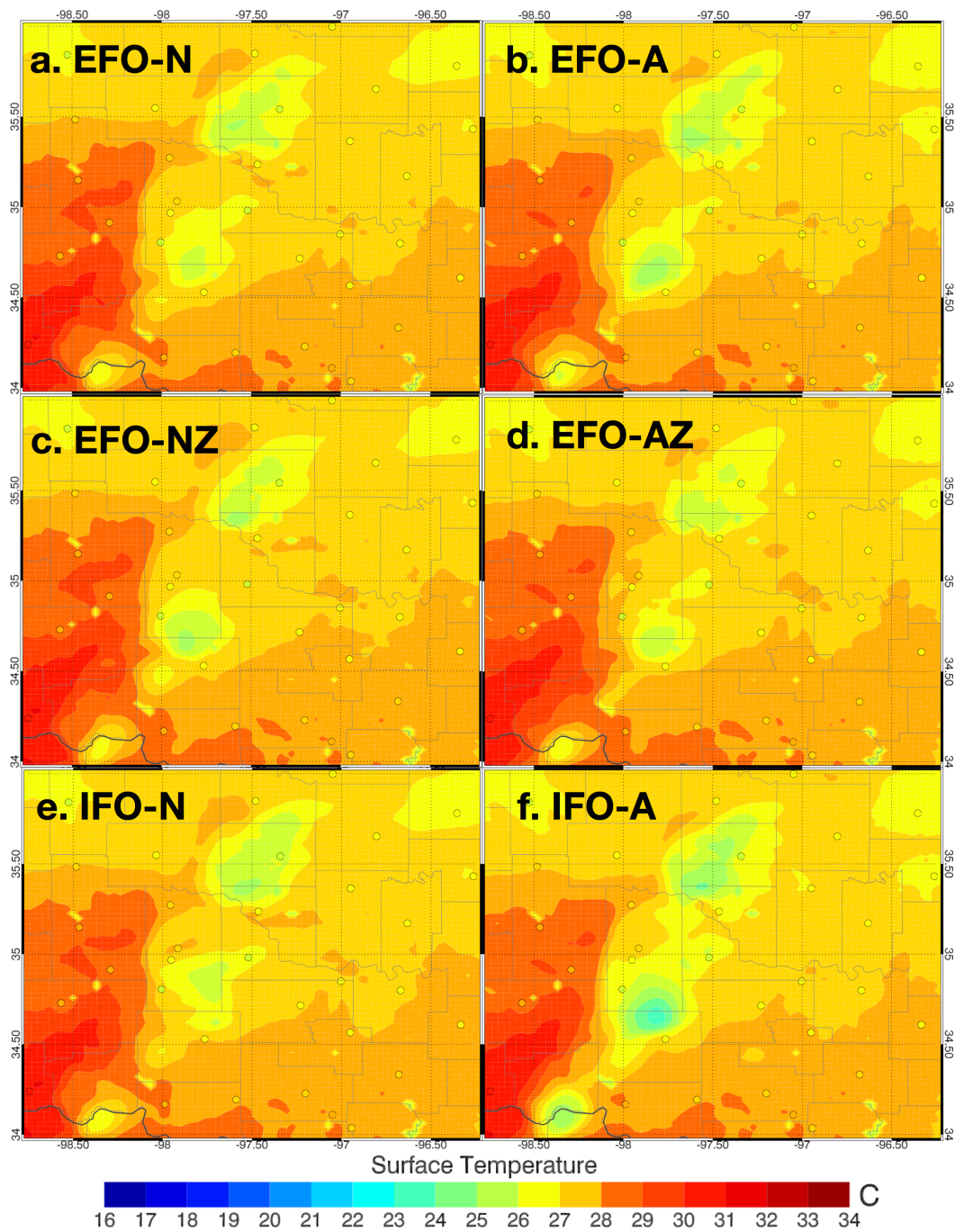
Figure 16. *Ensemble mean surface temperature at 2000 UTC 20 May 2013 for each experiment. Corresponding OK Mesonet observations are overlaid.*

Experiment:	Temp RMSE (°C)
EFO-N	1.98
EFO-A	1.86
EFO-NZ	2.10
EFO-AZ	2.02
IFO-N	2.04
IFO-A	1.70

Table 3. *Surface temperature RMSE between the ensemble mean 2000 UTC analysis from each experiment and OK Mesonet observations calculated for the 33 sites within domain plotted in Figure 16.*

d. Precipitation forecasts

The ensemble mean one-hour (2000 – 2100 UTC) total accumulated precipitation forecasts for each experiment were created and compared against the Stage 4 precipitation product available from the National Centers for Environmental Prediction for the same time period (Fig. 17). The Stage 4 precipitation analysis shows maxima of 1 and 3 cm accumulated precipitation during this 1 hour period for all of the analyzed storms. The greatest rainfall corresponds with storm #3 in southern OK. All experiments generate three swaths of precipitation corresponding to the three primary storms. However, the characteristics of the swaths differ between experiments. EFO-N and EFO-NZ precipitation forecasts associated with storm #1 do not compare favorably to observations, with forecast rainfall being too weak and displaced north and east. EFO-A and EFO-AZ generate slightly better placement, but still under-forecast total precipitation with storm #1. All EFO experiments perform well with storm #2, and significantly under forecast precipitation associated with storm #3. Finally, IFO-A generates the greatest amount of precipitation with the maximum precipitation cores agreeing well in magnitude and location to observations. ETS scores for 1 hour precipitation using thresholds of 0.5, 0.75, 1.0, 1.5, and 2.0 cm were calculated in a manner similar to that used for the idealized experiments. For all experiments, ETS decreases as threshold increases indicating poorer forecast skill for higher rainrates (Table 4). The non-AI experiments have little skill at thresholds > 1.5 cm, consistent with the qualitative assessment above. EFO-A and IFO-A perform better at all thresholds and maintain at least some skill up to 2.0 cm. In particular, the ETS score from IFO-A using the 2.0 cm threshold, 0.132, is larger than the ETS using the 1.5 cm threshold from the other five experiments.
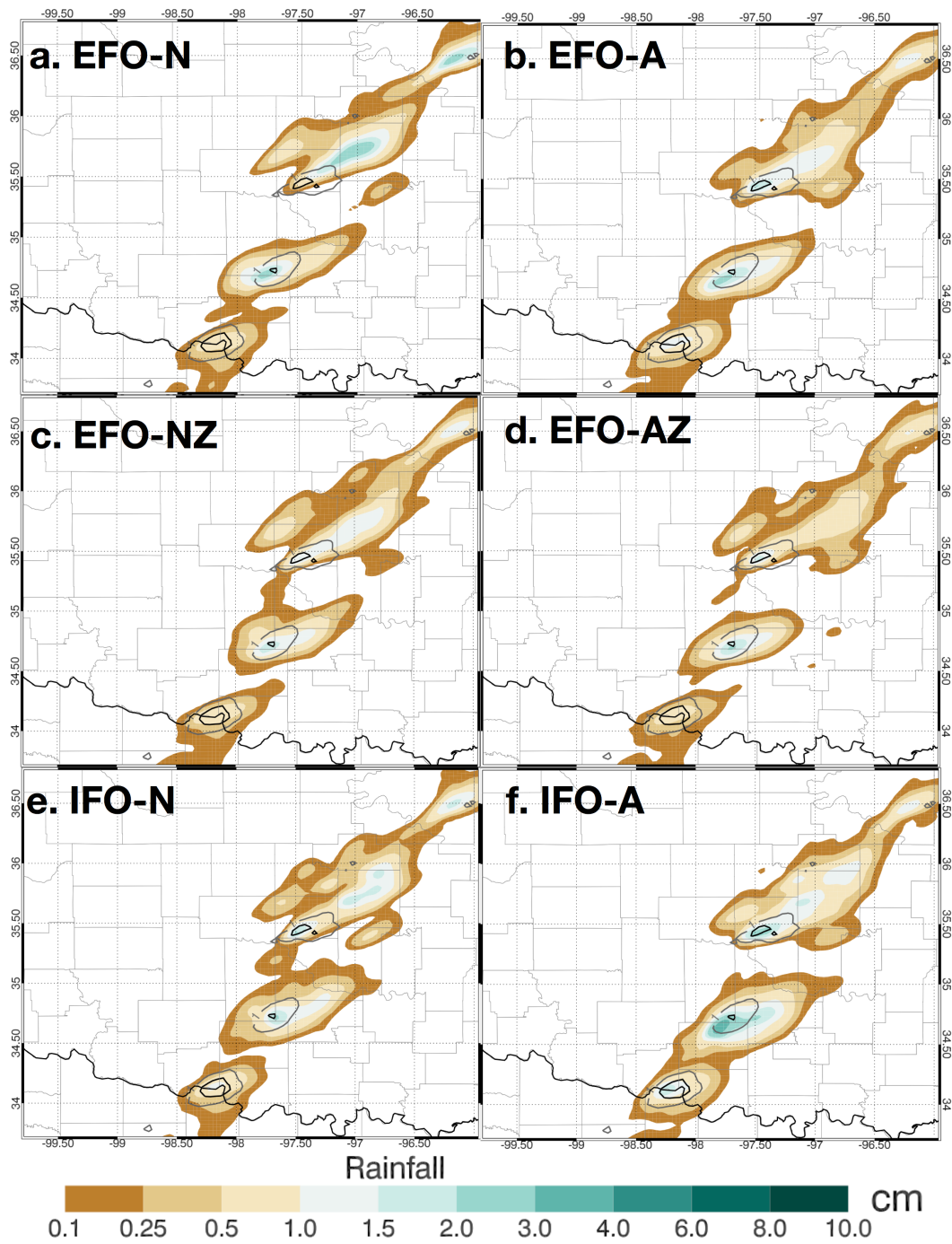
Figure 17. *Ensemble mean one-hour accumulated precipitation forecasts for each experiment between 2000 – 2100 UTC. Contours indicate location of Stage 4 precipitation observations greater than 1 and 3 cm over the same time period.*

Experiment:	0.5 cm	0.75 cm	1.0 cm	1.5 cm	2.0 cm
EFO-N	0.215	0.141	0.114	0.070	0.028
EFO-A	0.343	0.311	0.264	0.128	0.036
EFO-NZ	0.230	0.193	0.150	0.038	0.000
EFO-AZ	0.212	0.176	0.155	0.065	0.013
IFO-N	0.259	0.178	0.116	0.058	0.010
IFO-A	0.410	0.347	0.297	0.235	0.132

Table 4. One hour 2000 – 2100 UTC accumulated precipitation ETS for various thresholds against Stage 4 accumulated precipitation analysis for the same period. ETS values are averaged from ETS calculated from each ensemble member.

7. Conclusions

This study has shown that the method used to generate prior model reflectivity in observation space can make significant differences in hydrometeor characteristics of convection in the model analysis. Applying adaptive inflation only acts to increase the differences between methods such that hydrometeor concentrations can differ up to 100% between one experiment to another. Differences have been clearly demonstrated in both idealized and real-case experiments performed in this research. For the idealized experiments, the difference between using internal vs. external forward operators and interpolating in Z or dBZ space was generally small, but non-zero. The application of adaptive inflation produces large differences compared to the non-adaptive inflation runs and also increases the magnitude of the differences between various reflectivity assimilation methods. The experiment with adaptive inflation using the internal reflectivity forward operator (IFO-A) generated greatest QRAIN and QGRAUP concentrations, the strongest cold pools and heaviest precipitation. Analysis of precipitation forecasts compared to the nature run showed that IFO-A was more skillful at predicting moderate to heavy precipitation in the first 30 minutes. Differences between the experiments decrease as a function of time and most are similar by 1 hour as the impact of the model analysis decreases.

Comparisons between reflectivity assimilation methods for the 20 May 2013 case closely follow the trends observed in the idealized experiments. Differences in the non-AI experiments are more apparent with IFO-N generally producing higher QRAIN and QGRAUP concentrations then either EFO-N or EFO-NZ. As in the idealized experiments, applying AI greatly increased the differences between reflectivity assimilation methods. IFO-A generates convection with highest QRAIN and QGRAUP concentrations and stronger cold pools compared to the other experiments. It also produces the most skillful 1-h precipitation forecasts by correctly generating areas of heavy precipitation that the other experiments failed to generate or misplaced.

While arguments for and against internal vs. external reflectivity forward operators can be made, it is clear that the choice will have significant impacts on how reflectivity is assimilated into a storm-scale model. In general, the internal method produced the "stronger" storm, which sometimes allowed it to persist in model forecasts longer after data assimilation has ceased. This advantage does come with the increased complexity of adding the necessary microphysics information to the forward operator, but no additional computational cost is incurred. Another motivation for using the internal method is the likely increase in the number of diagnostic

variables to be assimilated into storm scale models in the near future. Examples include polarimetric radar observations, high-resolution satellite water vapor radiances, and satellite cloud property retrievals. Increasing the state vector to accommodate all these variables would not be practical and may also require very large computational memory since the entire state is stored. There is also the issue of mixing and matching internal vs. external forward operators for difference observation types. Since all these observations are directly related to hydrometeor variables, using a consistent approach with forward operators would likely produce more consistent results.

Acknowledgments

References

Aksoy, A., D. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part 1: Storm-scale analyses. Mon. Wea. Rev., 137, 1805-1824.

Aksoy, A., D. Dowell, and, C. Snyder, 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. Mon. Wea. Rev., 138, 1273–1292.

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus, 59A, 210-224.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. Tellus, 61A, 72-83.

——, and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. J. Atmos. Oceanic Technol., 24, 1452–1463.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. Bull. Amer. Meteor. Soc., 90, 1283–1296.

Crum, T. D., R. L. Alberty, 1993: The WSR-88D and the WSR-88D Operational Support Facility. Bull. Amer. Meteor. Soc., 74, 1669–1687.

Dawson II, D. T., L. J. Wicker, E. R. Mansell, R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. Mon. Wea. Rev., 140, 696–716.

Doviak, R. J., and D. S. Zrnic, 1993: Doppler Radar and Weather Observations. Dover, 562 pp.

Dowell, D., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. Mon. Wea. Rev., 132, 1982–2005.

Dowell, D. C., L. J. Wicker, C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale Analyses. Mon. Wea. Rev., 139, 272–294.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm- scale ensemble forecasting and data assimilation. J. Atmos. Oceanic Technol., 26, 911–927.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. Quart. J. Royal Meteor. Soc., 125, 723-757.

Gao, J., M. Xue, A. Shapiro, and K. K. Droegemeier, 1999: A variational method for the analysis of three-dimensional wind fields from two Doppler radars. Mon. Wea. Rev., 127, 2128–2142

Gao, J. and David J. Stensrud, 2012: Assimilation of Reflectivity Data in a Convective-Scale, Cycled 3DVAR Framework with Hydrometeor Classification. J. Atmos. Sci., 69, 1054–1065.

Hu, M. and M. Xue, 2007a: Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case. Mon. Wea. Rev., 135, 507–525.

Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, C. Karstens, 2015: A Comparison of Multiscale GSI-Based EnKF and 3DVar Data Assimilation Using Radar and Conventional Observations for Midlatitude Convective-Scale Precipitation Forecasts. Mon. Wea. Rev, 143, 3087–3108.

Jones, T. A., D. J. Stensrud, P. Minnis, and R. Palikonda, 2013a: Evaluation of a forward operator to assimilate cloud water path into WRF-DART. Mon. Wea. Rev., 141, 2272–2289.

Jones, T. A., J. Otkin, D. J. Stensrud, and K. Knopfmeier, 2013b: Assimilation of satellite infrared radiances and Doppler radar observations during a cool season Observing System Simulation Experiment. Mon. Wea. Rev. 141, 3273–3299.

Jones, T. A., J. Otkin, D. J. Stensrud, and K. Knopfmeier, 2014: Forecast evaluation of an Observing System Simulation Experiment assimilating both radar and satellite data. Mon. Wea. Rev. 142, 107-124.

Jones, T. A., D. J. Stensrud, L. Wicker, P. Minnis, and R. Palikonda, 2015: Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. Mon. Wea. Rev. 143, 165-194.

Jones, T. A. and D. J. Stensrud, 2015: Assimilating cloud water path as a function of model cloud microphysics in an idealized simulation. Mon. Wea. Rev. 143, 2052–2081.

Jones, T. A., K. Knopfmeier, D. Wheatley, G. Creager, P. Minnis, and R. Palikondo, 2016: The NSSL Multiscale Ensemble. Part 2: Combined radar and satellite assimilation. Wea. Forecasting, 31, 297-327.

Jung, Y., M. Xue, M. Tong, 2012: Ensemble Kalman Filter Analyses of the 29–30 May 2004 Oklahoma Tornadic Thunderstorm Using One- and Two-Moment Bulk Microphysics Schemes, with Verification against Polarimetric Radar Data. Mon. Wea. Rev., 140, 1457–1475.

Kalman, R. E., 1960: A new approach to linear filtering and pre- diction problems. J. Basic Eng., 82, 35–45.

Majcen, M., P. Markowski, Y. Richardson, D. Dowell, and J. Wurman, 2008: Multipass objective analyses of Doppler ra- dar data. J. Atmos. Oceanic Technol., 25, 1845–1858.

Rinehart, R. E., 1997: Radar for Meteorologists, Third Edition. pp 428.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X-Y. Huang, W. Wang, and J. G. Powers, 2008: A description of the Advanced Research WRF version 3. NCAR Tech Note NCAR/TN- 475+STR, 113 pp. [Available from UCAR Communications, P.O. Box 3000, Boulder, CO 80307.]

Stensrud, D.J, J-W. Bao, and T. T. Warner, 2000: Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems. Mon. Wea. Rev., 128, 2077–2107.

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk micro- physics scheme. Part I: Description and sensitivity analysis. Mon. Wea. Rev., 132, 519–542.

——, P. R. Field, R. M. Rasmussen, and W. R. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. Mon. Wea. Rev., 136, 5095–5115.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. Mon. Wea. Rev., 133, 1789–1807.

Tong, M. and M. Xue, 2008a: Simultaneous Estimation of Microphysical Parameters and Atmospheric State with Simulated Radar Data and Ensemble Square Root Kalman Filter. Part I: Sensitivity Analysis and Parameter Identifiability. Mon. Wea. Rev., 136, 1630–1648.

Tong, M. and M. Xue, 2008b: Simultaneous Estimation of Microphysical Parameters and Atmospheric State with Simulated Radar Data and Ensemble Square Root Kalman Filter. Part II: Parameter Estimation Experiments. Mon. Wea. Rev., 136, 1649–1668.

Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. Mon. Wea. Rev., 110, 504–520.

Wheatley, D. M., K. Knopfmeier, T. A. Jones, G. Creager, 2015: Storm-scale Data Assimilation and Ensemble Forecasting with the NSSL Experimental Warn-on Forecast System. Part I: Radar Data Experiments., Wea. and Forecasting. 30, 1795–1817.

Whitaker, J. S., T. M. Hamill, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. Mon. Wea. Rev., 136, 463–482.

Wilks, D. S., 2006: Statistical Methods in the Atmospheric Sciences. 2nd ed. Academic Press, 627 pp.

Yussouf, N. and D. J. Stensrud, 2010: Impact of phased-array radar observations over a short assimilation period: Observing system simulation experiments using an ensemble Kalman filter. Mon. Wea. Rev., 138, 517–538.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, D. J. Stensrud, 2013: The ensemble Kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. Mon. Wea. Rev., 141, 3388–3412.